Gait Recognition as a Service for Unobtrusive User Identification in Smart Spaces

CHENGWEN LUO, JIAWEI WU, JIANQIANG LI*, and JIA WANG, the College of Computer Science and Software Engineering, Shenzhen University, China
WEITAO XU, Department of Computer Science, City University of Hong Kong, Hong Kong, China
ZHONG MING*, the College of Computer Science and Software Engineering, Shenzhen University, China
BO WEI, the Department of Computer and Information Sciences at Northumbria University, England
WEI LI and ALBERT Y. ZOMAYA, The University of Sydney, New south wales, Australia

Recently Internet of Things (IoT) has raised as an important research area that combines the environmental sensing and machine learning capabilities to flourish the concept of smart spaces, in which intelligent and customized services can be provided to users in a smart manner. In smart spaces, one fundamental service that needs to be provided is accurate and unobtrusive user identification. In this work, to address this challenge we propose a Gait Recognition as a Service (GRaaS) model, which is an instantiation of the traditional Sensing as a Service (S^2aaS) model, and is specially deigned for user identification using gait in smart spaces. To illustrate the idea, a Radio Frequency Identification (RFID) based gait recognition service is designed and implemented following the GRaaS concept. Novel tag selection algorithms and attention-based Long Short-Term Memory (At-LSTM) models are designed to realize the device layer and edge layer, and achieves a robust recognition with 96.3% accuracy. Extensive evaluations are provided, which show that the proposed service has accurate and robust performance and has great potential to support future smart space applications.

Additional Key Words and Phrases: IoT, RFID, Gait Recognition, User Identification, Attention-based LSTM

ACM Reference Format:
Chengwen Luo, Jiawei Wu, Jianqiang Li*, Jia Wang, Weitao Xu, Zhong Ming*, Bo Wei, Wei Li, and Albert Y. Zomaya. 2019. Gait Recognition as a Service for Unobtrusive User Identification in Smart Spaces. 1, 1 (November 2019), 21 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

In the past decade, the Internet of Things (IoT) rises as an emerging computing paradigm that interconnects distributed sensing devices such as sensor nodes, smartphones, wearables, etc., through various networking technologies to enable smart services that empower different applications

[1]* Corresponding author

Authors' addresses: Chengwen Luo, chengwen@szu.edu.cn; Jiawei Wu, gary036@163.com; Jianqiang Li*, lijq@szu.edu.cn; Jia Wang, jia.wang@szu.edu.cn, the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, 518060; Weitao Xu, weitaoxu@cityu.edu.hk, Department of Computer Science, City University of Hong Kong, Hong Kong, Hong Kong, China; Zhong Ming*, mingz.xu@szu.edu.cn, the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, 518060; Bo Wei, the Department of Computer and Information Sciences at Northumbria University, Newcastle, England, bo.wei@northumbria.ac.uk; Wei Li, weiwilson.li@sydney.edu.au; Albert Y. Zomaya, albert.zomaya@sydney.edu.au, The University of Sydney, Sydney, New south wales, Australia.

XXXX-XXXX/2019/11-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

in smart health [1, 2, 8, 15, 20, 27], agriculture, city management, industries, and so on. As the advances in the sensing, communication and data processing technologies [5, 6], IoT allows people to be connected anytime, anywhere, with anything and anyone [22]. Enabled by the vision of IoT, the way people interact with the surrounding environments will also be significantly changed. With the sensing capability, the environment can perceive and react to people without requiring them to equip special-purpose devices. This prospers the notion of *smart spaces*[33], in which intelligent and customized services can be provided *in situ* enabled by the context-awareness achieved through the sensing technologies. Applications in the domain of smart home, smart building, smart cities, and even smart nations can then be realized by utilizing the sensing, communication, and data processing capability with the IoT infrastructure.

Unobtrusive User Indetification. In smart spaces, one fundamental problem is to achieve unobtrusive user identifications in order to provide customized services to users. For example, in smart home applications, customized music list can be played once one particular family member returns home, and different light settings can be triggered based on the identity of different family members. User identification is also critical in the security management in smart spaces[10]. For example, whether to authenticate one user to enter the space should also depend on the reliable and accurate the user identtity detection. The obtrusiveness of identification process should be minimized to improve the user experience. Therefore, in the design of smart spaces, accurate and unobtrusive user identification is one fundamental service that needs to be realized to enable the intelligent authentication, customization, and improve user experiences inside the space.

Much existing work has studied the user identification problem using different approaches such as video processing [11], acoustic processing [24], biometrics-based recognition such as face recognition[9], fingerprint recognition[4], and iris recognition[18]. While these work has successfully demonstrated the feasibility of accurate user identification, they have their relative limitations in the application domain of smart space. For example, video-based, acoustic-based, and face recognition approaches are usually subject to high computational cost and privacy concerns. Fingerprint-based and iris-based recognition often require users to perform particular actions to get identified or authenticated, hence are obtrusive to use and are unable to provide continuous identifications. Gait recognition is an emerging technique [38] that exploits the unique walking pattern of user as the biometric signal for identifications. Using gait for identification has several advantages. First, the gait can be opportunistically captured for identification without explicitly asking user to perform certain actions. For example, gait recognition can be performed at the entrance of the space so that the user gets authenticated automatically when enters the space. Second, some gait recognition approaches such as the wireless-based gait recognition does not require to record videos, images, or sounds, hence minimizes the privacy concerns. Third, gait based processing are often less computationally intensive, therefore reducing the computation costs. With these benefits, we envision gait-based recognition approach will continue to gain its popularity in future smart spaces and become one important service for user identification.

Gait-based approaches are first investigated using cameras [19, 29, 30]. However, as discussed above such approaches are often subject to high privacy concerns and are not always suitable for use in smart spaces. Recently due to the rise of sensor-equipped devices, researchers have proposed the motion sensor based approaches to achieve gait recognition for device user identification and secure communication [36]. However, such approaches require users to hold or wear the devices during walking, which makes the system infeasible for identifying multiple users. To achieve less obtrusive gait-based identification, recently researchers started to study the wireless-based gait recognition. Wireless signals such as WiFi are used to extract features and recognize the user when he/she is walking by [38]. These approaches have opened up new opportunities for low-cost, accurate, and unobtrusive user identification in smart spaces. *In this paper, our focus therefore is to* *design the service architecture of wireless RFID based gait recognition for efficient user identification in smart spaces.*

Several challenges exist to design an efficient service architecture for gait-based identification in smart spaces. First, to support pervasive IoT scenario, the resource constraint of IoT devices should be met. Since most IoT devices have strict resource constraints, the computational cost should be kept as small as possible to reduce the energy consumption used in gait recognition. Second, in real-world environment, user identification can be critical since it might closely related to security applications, e.g., gait-based authentication services. Therefore accuracy is an important concern since otherwise the security cannot be guaranteed. Third, in smart spaces, the user might change frequently and the environment dynamics will inevitably affect the performance of user identification, and the design of the identification services should take all these factors into considerations.

Contributions. In this paper, we therefore propose the gait recognition as a service (GRaaS) for unobtrusive user identification in smart spaces. To accommodate the resource constraint and timely response requirement of the IoT applications, we design the edge-cloud architecture for service distribution. To achieve the cost effectiveness and unobtrusiveness, a RFID-based solution is proposed for wireless-based gait recognition. A novel tag selection algorithm and attention-based Long Short-Term Memory (LSTM) model are used for gait recognition to achieve the high accuracy and robustness in dynamic environments. By integrating all above functionalities, an efficient wireless-based gait-recognition service architecture is designed for unobtrusive and accurate user identification in smart spaces. In summary, we made the following contributions:

- We design and implement the edge-cloud service architecture for efficient service management to achieve the high responsiveness and low computational overhead for IoT devices.
- We adopt the RFID-based wireless sensing in the sensing layer to achieve the unobtrusiveness for gait-based user identification. To the best of our knowledge, our system is the first in the literature that achieves accurate gait recognition using RFID signals.
- We design a novel tag selection algorithm and attention-based LSTM model for accurate and robust identification in dynamic environments. Extensive evaluations show that current design achieves high accuracy in real-world environments and low computational cost for IoT devices.

The rest of the paper is organized as follows. First, Section 2 discusses the related work. Then Section 3 provides the overall architecture for the gait recognition service management. Section 4 provides our detailed design for the RFID-based gait recognition service. Followed by that, we evaluate the overall system in Section 5, discuss the limitation and future work in Section 6, and finally conclude the paper in Section 7.

2 RELATED WORK

Our work is closely related to several categories of work in the literature. In this section, we mainly discuss the recent developments in the related *Sensing as a service (S^2aaS)*, *RF-based gesture recognition*, *RF-based activity recognition*, and *WiFi-based gait recognition*.

Sensing as a Service (S^2aaS). As the rise of cloud computing, Everything as a Service (XaaS) [3] a category of models introduced to provide resources as a service based on the cloud platform [22]. Among them, Sensing as a Service model is proposed as a solution based on the Internet of Things infrastructure. Four conceptual layers are designed in the model: sensor data owners, sensor data publishers, extended service providers, and sensor data consumers. The S^2aaS provides a model for sensor data exchange for IoT applications. As a special case, the GRaaS model proposed in this work is a realization of the S^2aaS model and speficially focuses on the gait recognition services for smart spaces.

RF-based Gesture Recognition. The Radio Frequency (RF) based gesture recognition has become popular recently due to its non-invasive nature, low cost, and universal usability. Using the RF signal such as RFID and WiFi, the gestures can be sensed based on the fact that different motion patterns will introduce different interruptions to the signal received. Most wireless gesture recognition systems are based on RFID or WiFi. For example, RFID-based gesture recognition GRFID[40] proposes a gesture recognition system using complex feature extraction techniques to recognize a few gesture types. Wisee[23] is a WiFi based gesture recognition system and successfully achieves high accuracy in the gesture recognition at the expense of complex system implementation and modified hardware.

RF-based Activity Recognition. Following the similar rationale, RF-based activity recognition [32] extends the gesture recognition capability to recognize general activities such as fall detection [37], walking activity tracking [35], etc. [37] uses the RSSI signal of RFID to identify different actions, and proposes a hidden Markov model (*HMM*) to infer the action changes. However, such RF-based activity recognition systems are prone to errors in dynamic environments and more sophisticated algorithms or system design are required to achieve higher accuracy and robustness in real-world environments.

WiFi-based Gait Recognition. Gait recognition based on WiFi has gradually become a popular research area recently. WiFi-ID [39] captures channel state information (CSI) of WiFi and extracts gait characteristics to achieve gait recognition. WiFi-ID captures complex time domain features, and use machine learning approaches to achieve accurate classification. The system cannot be extended to a large group and only limits to a small group of less than six people. WiFiU system[31] leverages commercial WiFi devices to capture fine-grained gait features to identify users. Although the system achieves the classification accuracy of over 80% for a large space and a group of 50 persons, the system requires large dataset for training, complex spectrum analysis and processing, resulting a high computational cost.

Our work differs from all above works in that we focus on the design of the Gait recognition as a service model for Internet of Things applications in smart spaces. We design and implement the edge-cloud architecture for user identification in smart spaces, and RFID-based gait recognition techniques to achieve an accurate and robust recognition in real-world environments.

3 SERVICE MANAGEMENT ARCHITECTURE

Recently, Sensing as a Service (S^2aaS) model has been proposed to support the vision of smart cities (SC) [22]. The S^2aaS is a vision that promotes data exchange between data owners and data consumers and consists of four roles, the *Sensor Data Owners*, *Sensor Data Publishers*, *Extended Service Providers*, and *Sensor Data Consumers*. The S^2aaS provides a generic architecture for data generated from IoT devices to be exchanged and to support various extended services such as data analytics for smart cities. While the S^2aaS architecture provides structural approaches for the data exchange and service management in IoT, it covers a broad range of applications and does not differentiate data type and focus on specific application scenarios.

As the gait recognition becomes one important functionality for intelligent user identification in smart spaces, we envision a Gait Recognition as a Service (GRaaS) architecture to support efficient user identification for applications in IoT. As depicted in Figure 1, the GRaaS model consists of four conceptual layers: the device layer, the edge layer, the service provider layer, and the application layer.

- **Device Layer.** This layer manages the sensors-equipped devices to collect gait-related signals for user identification. For example, a camera can provide surveillance to a certain area and collect vision-based data for gait recognition, and the commodity WiFi access point can be

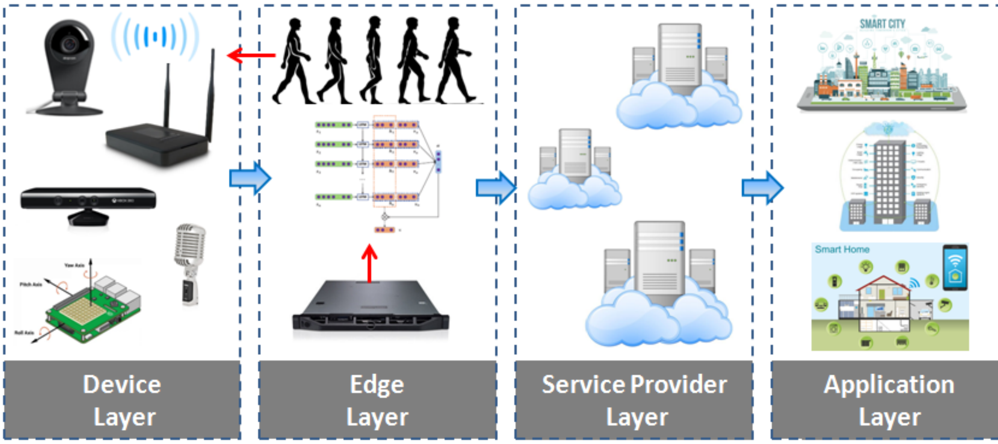


Fig. 1. The Gait Recognition as a Service Model

the device to collect WiFi CSI signals for gait recognition [39], etc. The devices in this layer normally should be within the vicinity of the area of interest in the smart space to provide the real-time sensor input for further processing. To improve the data quality and reduce the data transmission overhead, data preprocessing techniques might be applied in this layer to reduce data size and eliminate noises.

- **Edge Layer.** The network edge lies between the user devices and the cloud server in the path of sensor data flow. This layer is usually optional but is designed to significantly improve the overall energy efficiency and timely response for IoT applications. Due to the resource constraint of IoT devices, computationally heavy tasks should be offloaded to the edge for better energy efficiency. And since edge is physically close to the user device, the edge layer provides faster responses than traditional cloud computing paradigm. In GRaaS, the machine learning algorithms which are usually computationally intensive should be offloaded to the edge.
- **Service Provider Layer.** This layer consists of service providers (SP) which provide cloud platforms for service exchange between the service generating edge and the customers. The edge runs gait recognition algorithms and generates user identification services, and the services are registered in the cloud-based service providers, which serve as central platform for networking, storage, and billing. Each SP can only have access to the services that are registered with it. The SPs acts as the central platform to coordinate and multiple services generating edges and the final applications that uses the gait recognition services.
- **Application Layer.** The application layer consists of customers who are interested in the gait recognition services. The customers might utilize the gait recognition for various purposes such as identification, authentication, tracking, etc. In order to get access to the service, applications are required to register with the SPs, and one application might register with multiple SPs in order to cover a certain area of interest. With the GRaaS architecture, application users can get access to the gait recognition services seamlessly to build up identification systems, authentication systems, etc.

4 RFID-BASED GAIT RECOGNITION SERVICE FOR GRAAS

In this section, we present a case study by designing and implementing a RFID-based gait recognition service following the GRaaS architecture. RFID is cheap and commonly available for IoT applications.

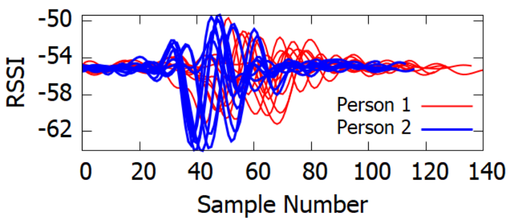


Fig. 2. Different RSSI patterns caused by different gaits

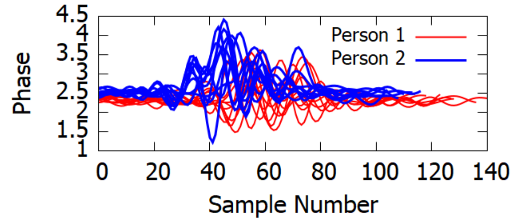


Fig. 3. Different phase profile caused by different gaits

Deploying RFID is cheap and is cost-effective even large number of tags are deployed. Besides, the wireless nature of RFID-based gait recognition minimizes the privacy concerns comparing with vision-based or acoustic-based identification approaches. RFID-based gait recognition also provides continuous identification and is less obtrusive to use without requiring users to explicitly perform certain actions. Due to these advantages, RFID-based gait recognition is a promising service for user identification in smart space applications. In the following subsections, we detail the design of the RFID-based Gait Recognition Service for GRaaS.

4.1 Technical Background

In this section, we briefly introduce the technical background of Ultra High Frequency (UHF) RFID systems and their signal propagation information that affects the design of gait recognition system. We are particularly interested in the RSSI and phase information, which is the fundamental information that is used for gait recognition.

4.1.1 RFID and Gait Recognition. In UHF RFID systems, the reader first transmits an electromagnetic wave to the tag through the antenna, and the tag extracts the energy required for the work from the electromagnetic wave. After the electromagnetic wave encounters the target such as the tag, a part of the electromagnetic wave is scattered back to the reader, and the reader receives the demodulated backscattered electromagnetic wave signal to obtain the data information of the tag. It is a two-way communication process: one link is the reader-to-tag signal, and the other link is the tag-to-reader signal. Only two links are successfully communicated, and the RF device completes a valid communication. Thus the actual identification distance of the system is determined by the shortest link communication path distance. Although the radio frequency signal used in RFID technology can penetrate the materials such as paper, wood and plastic, it is still affected by many factors, such as metal, liquid and so on. Among them, human body also has a significant impact on the radio frequency signals. Most of the radio frequency signals will be absorbed by the human body when passing through the human body, so the actual radio frequency signals received will have observable changes. Therefore, different body shapes and different walking patterns will introduce marked changes in the interruptions to the signal propagation between the reader and the tag. This becomes the fundamental rationale to classify different gaits and to identify different people using RFID signals.

4.1.2 Key Gait-related RFID Signal Information. Typically, signal information such Received Signal Strength Indicator (RSSI) and phase are two key information that can be used to differential different gaits when users are walking by between the tags and the reader. Here we briefly introduce RSSI and Phase information.

- **RSSI**. RSSI is the difference between the true received signal strength and the optimal received signal strength level, and its implementation is carried out after the baseband receiving filter of the reverse receiving channel. In the passive RFID system readers, RSSI mainly refers to the signal strength of the tag reflection signal. Currently, the RSSI value is used to determine the distance from the reader. The larger the RSSI value, the closer the tag is to the reader. That is, the RSSI value is inversely proportional to the distance between the tag and the reader and is normally modeled with the log-distance path loss (LDPL) model [7]:

$$RSSI = p_0 - 10\gamma log\Delta d + \varepsilon \tag{1}$$

where p_0 is the RSSI at a distance of one meter, γ is the rate of fall, Δd is the distance between the reader and the tag, and ε is a random variable to capture the variations of the RSSI measurements. Due to the multi-path effect, RSSI is quite complicated in practical scenarios, especially for passive tags. Human body has significant influence on RF signals, most of the radio frequency signals will be absorbed by human body after passing through human body, as well as noise interference in the environment, resulting the actual received RSSI to produce marked changes. Figure 2 shows the RSSI sequences captured when asking two users to walk for 10 times independently. As shown in Figure 2, since each person's gait is unique, the RSSI changes are different when different users are walking by, making this information a suitable input for gait recognition.
- **Phase**: Phase is the basic property of RF signal in the physical layer. The phase value of an RF signal describes the degree of deviation between the receiving signal and the sending signal, ranging from 0 to 2π. Let Δd describes the distances between the reader and the RFID tag, then the phase measurement θ can be computed as [25]:

$$\Theta = (2\pi \frac{2\Delta d}{\lambda} + \mu) \mod 2\pi \tag{2}$$

where λ is the wavelength and μ is system noise. Within the signal coverage range, as the human body absorbs most of the RF signal, the actual received RF phase value will change significantly. Due to the uniqueness of the human gait, the signal changes in the phase profile is also unique, therefore using the phase values captured by the RFID reader can also be for gait recognition. Figure 3 shows the phase profile captured when two users walk for 10 times each. As shown in Figure 3, a marked difference can be observed, which illustrates the feasibility of phase-based gait recognition.

4.2 Design Overview

Following the GRaaS architecture, we design the RFID-based gait recognition service for user identification in smart spaces deployed using commercial off-the-shelf RFID devices. The design overview of the proposed system is shown in Figure 4. The whole system consists of four layers. The device layer consists of two parts, transmitting and receiving RFID radio frequency signal antenna as reader and passive RFID tags as transponders. A tag selection algorithm is first performed to identify the optimal combination of tags that should be used in the processing to improve the accuracy of the recognition. The RFID reader extracts RSSI signals and phase profiles as input for further processing. Due to the limited computational capability of the RFID reader, the extracted data is uploaded to a local edge device for processing. The edge device receives the RSSI and phase profile and performs gait recognition algorithms for user identification. An attention-based LSTM model is then used for training and classification of the input data to achieve an accurate recognition at the edge layer. The service is registered with the service provider layer, which provides a cloud

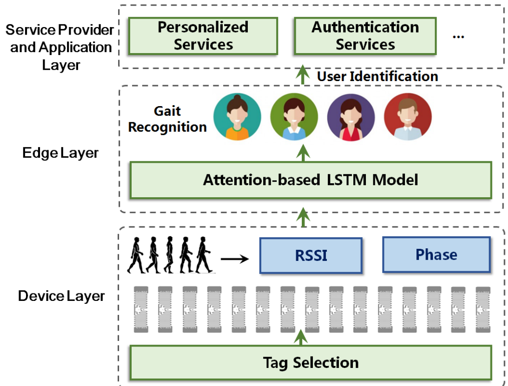


Fig. 4. Overview of RFID-based gait recognition service for user identification in smart spaces

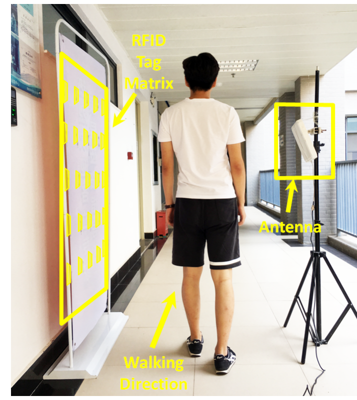


Fig. 5. System deployment and operational scenario

platform for customers to access the server from anywhere. And applications such as authentication in smart spaces then can be built in the application layer.

The working scenario is described as follows: The devices are deployed within the area of interest, typically an ordinary corridor, with one side covered with RFID tags, and the other side the RFID antenna and reader to capture the RFID signals when users are walking by. An RFID antenna continuously emits radio-frequency signals, and the tag receives the radio-frequency signals from the antenna. The antenna receives and reads the decoded information and sends it to the edge device for data processing. An edge device can be a local server or simply a router with programming capability. When a person walks along the corridor, his/her gait affects the signal propagation in a unique way, resulting in the changes in the RSSI and phase values. As a result, when someone passes through the corridors, the multi-path of each tag's signal propagation changes accordingly, which can be captured by the RSSI and phase information and used for gait-based user identification. After receiving the RSSI and phase profiles, the edge first pre-processes the raw sensing data to filter and normalize the raw data. Next, the edge applies the attention-based LSTM model to train and classify the preprocessed data to effectively learn the characteristics of gait and to achieve accurate gait classification. To improve accuracy, tag selection process is repeatly performed to select the optimal subset of tags for sensing and classification.

4.3 Device Layer: Tag Deployment and Selection

In the device layer, tags are deployed to continuously collect the RSSI and phase signals for sensing different gait patterns. In this section, we first introduce tag deployment schemes and then tag selection algorithms for improving the efficiency of gait recognition.

4.3.1 Tag Deployment. In the device layer, to achieve gait recognition the first important challenge is to design the tag deployment schemes for efficient gait recognition. Due to the complicated nature of radio signals in the indoor environment, different configuration of RFID tags determines the unique combination of RSSI and phase profiles, so as to directly affect the recognition performance. For example, if two tags are closely placed within a certain distance, signal collision problems will occur [28] and affects the final data collection for gait sensing. The number of tags used in our system has also an impact on the final performance of gait recognition. As a result, multiple tags are deployed in the system combined with an efficient tag selection algorithm to determine the optimal subset of tags to used for gait recognition.

The sampling rate of a typical UHF RFID reader ranges from 25 Hz to 30 Hz. As a single tag contributes only single channel input, the data size is insufficient for gait recognition purpose. On the other hand, the reading of a single tag can be easily affected by the environmental noises, making recognition results less reliable if only one tag is used. As a result, it is undesirable to use a single tag for gait recognition, therefore instead in our system a multi-tag solution is adopted.

The other important concern is the configuration of multiple tags. One possible configuration is to use a single-line topology and put all tags in a single line. However, although a single-line tag configuration can capture signal changes properly, it might not be able to detect fine-grained body motions and is less sensitive to gait movements. The reason is that the human gait movements generally consist of whole-body movements, which involve leg, torso, arm, head, etc. Therefore using a single-line configuration cannot effectively capture all gait movements.

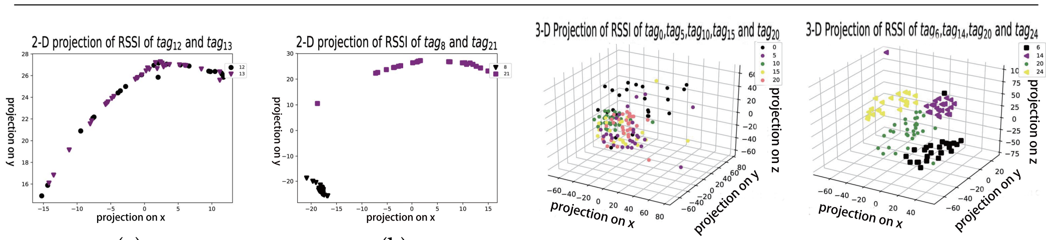
To achieve higher precision and finer-grained recognition accuracy, a tag matrix is a better configuration to capture signals affected by the different body movements of the gait. For example, the lower half of the tag matrix mostly corresponds to the action of the torso and leg; and the upper half of the tag array mostly corresponds to the movement of the upper body of the person, such as shaking the head and waving the arm. In this way, we can use the RSSI and phase profiles collected to achieve a more robust gait recognition. The final tag deployed is illustrated in Figure 5. As shown in the figure, 5×5 RFID tag matrix is deployed for gait recognition. To achieve better performance, a novel tag selection algorithm is proposed to compute the optimal combination of tags in the matrix.

4.3.2 Tag Selection. Tag selection is used to identify the optimal subset of tags to use in order to reduce computation, improve accuracy, and increase cost effectiveness. Similar studies have shown that optimal tag configuration can reduce the error caused by the fluctuation of RFID signals [12]. The key idea of tag selection is to analyze the similarities between the signals collected from different tags, identify tags with high correlations and those that can represent the whole tags, and use this subset as the selected tags for gait recognition. The motivating example is illustrated in Figure 6. As shown in Figure 6 (a), when a user walks pass the corridor deployed with reader and tags, the collected RSSI values of tag_{12} and tag_{13} are highly similar, indicating that the correlation between RSSI signals of both tags is high. Since they are placed on the same row in the tag matrix, the result indicates that one of them is redundant and can be represented by the other. Figure 6 (b) shows another example comparing the RSSI signals of tag_8 and tag_{21}. We can see that little correlation exists in their RSSI signals and these tags captures different properties of gait movements. In this case both tags cannot represent each other. Figure 6 (c) shows the 3D projection of RSSI values of $tag_6, tag_{14}, tag_{20}, tag_{24}$ and tag_{12}. RSSI of these tags are clearly distinguishable, however the 3D projection of RSSI values of $tag_0, tag_5, tag_{10}, tag_{20}$ and tag_{20} indicates that they are are highly correlated as shown in Figure 6 (d), and redundant information can be removed for better recognition performance. To eliminate redundancy and find the optimal subset of tags, we used three algorithms for tag selection: *Relief F*, *F-Statistics* and *Deep Reinforcement Learning on Tag Selection*. Before introducing the detailed tag selection algorithms, for the sake of convenience we summarise the notations in Table 1, which will be used in our later algorithm descriptions.

RSSI and phase values received are expressed as follows: Let $R \in \mathbb{R}^d$ and $P \in \mathbb{R}^d$ (d is the number of tags) be the domain of observable RSSI r and phase p. And $L \in \{1, ..., K\} \subset \mathbb{R}$ be the domain of output subject label l. Suppose we have n RSSI, phase and subject label pairs $\{(r_i, l_i)|r_i \in R, l_i \in L, i = 1, ..., n\}$, $\{(p_i, l_i)|p_i \in R, l_i \in L, i = 1, ..., n\}$ and $\{(r_i, p_i, l_i)|r_i \in R, p_i \in R, l_i \in L, i = 1, ..., n\}$.

Table 1. Summary of notations

Notations	Meaning
D	Dataset
m	the number of sampling
k	the number of nearest neighbor samples
N	the number of tags
R	randomly select a sample R from dataset D
$H_j(j = 1, .., k)$	the k nearest neighbors H_j from the same (the same tag) sample of R
$M_j(C)(C \notin class(R))$	the k nearest neighbors $M_j(C)$ from each different class(the different tag) sample set
n	the number of people
n_j	n_j is the number of samples in jth subject
f_i^j	f_i^j is the mean value of tag_i in the jth subject.
f_i	f_i denotes the mean value of tag_i in the dataset.
$f_{k,i}^j$	$f_{k,i}^j$ denotes the mean value of kth samples of tag_i in j_th subject
S	States
A	Actions
γ_i	the upper boundary
η_i	the lower boundary
$M_i^{'}$	the tag adjustments
P_n	the accuracy of pre-defined model
r	Rewards
$r^{'}$	the difference in the accuracy between the $(n-1)$th and nth iteration
σ	a strong stimulation
v_t	a normalized discounted reward
β	the learning rate of TSNet
θ	the parameter of TSNet



(a) (b) (c) (d)

Fig. 6. The correlation between tags

The training dataset can be described as:

$$
\begin{aligned}
R &= [r_1, ..., r_n] \in \mathbb{R}^{d\times n} \\
P &= [p_1, ..., p_n] \in \mathbb{R}^{d\times n} \\
L &= [l_1, ..., l_n] \in \mathbb{R}^{n}
\end{aligned} \tag{3}
$$

With these notations, we are now ready to discuss the detailed tag selection algorithms.

- **Relief F**. Relief F is a feature weighting algorithm, which assigns different weights according to the correlation of each feature and category. Features whose weight are less than a certain threshold will be removed. The correlation between features and categories in the algorithm is based on the ability of features to distinguish close-range samples. It calculates the weight of each tag and finally obtains the average weight of each feature. The greater the weight of the feature, the stronger the classification ability of the feature, otherwise the weaker the ability to classify using the feature. The details of Relief F are outlined in Algorithm 1. The weight of each tag_i is updated by the following evaluation function:

$$W(i) = W(i) - \sum_{j=1}^{k} diff(i, R, H_j)/(mk) + \sum_{C \notin class(R)} [\frac{p(C)}{1 - p(class(R))} \sum_{j=1}^{k} diff(i, R, M_j(C))/(mk)] \tag{4}$$

where dataset D can be RSSI, phase or a combination of RSSI and phase. $M_j(C)(C \notin (class(R))$ represents the j_{th} nearest neighbor sample and $diff(i, R_1, R_2)$ represents the difference between R_1 and R_2 on tag_i, as follows:

$$diff(i, R_1, R_2) = \begin{cases} \frac{|R_1[i] - R_2[i]|}{max(i) - min(i)}, \text{if } n \text{ is continuous} \\ \begin{cases} 0, & \text{if } n \text{ is discrete and } R_1[i] = R_2[i] \\ 1, & \text{if } n \text{ is discrete and } R_1[i] \neq R_1[i] \end{cases} \end{cases} \tag{5}$$

Algorithm 1: Relief F

Input: Dataset D, the number of sampling m, the number of nearest neighbor samples k, the number of tags N

Output: The Weight of each tag T

```
Set all tag weights to 0, and set T to an empty set;
for i = 1, 2, ..., m do
    We randomly select a sample R from dataset D;
    Find the k nearest neighbors H_j(j = 1, .., k) from the same (the same tag) sample of R,
      and find k nearest neighbors M_j(C)(C ∉ class(R)) from each different class (the
      different tag) sample set;
end
for i = 1, 2, ..., N do
    Execute equation (4);
end
```

- **F-Statistics**. An F-statistics is a value when running an ANOVA (analysis of variance) test or a regression analysis to determine if the means between two populations are significantly different. An F-score is computed to decide whether a group of variables are jointly significant. The F-score is calculated using the following formula:

$$F_i = \frac{\sum_{j=1}^{n}(\bar{\theta_i^j} - \bar{\theta}_i)^2}{\sum_{j=1}^{n} \frac{1}{n_j - 1} \sum_{k=1}^{n_j}(\theta_{k,i}^{j} - \bar{\theta_i^j})^2} \tag{6}$$

where θ can be RSSI, phase or a combination of RSSI and phase, n is the number of people. n_j is the number of samples in jth subject. $\bar{\theta}_i^j$ is the mean value of tag_i in the jth subject. $\bar{\theta}_i$

denotes the mean value of tag_i in the dataset. And $\theta^j_{k,i}$ denotes the mean value of k_{th} samples of tag_i in j_{th} subject. The numerator represents the difference between positive and negative sets, and the denominator represents the distinction between each of the two groups. The greater the F-score, the more likely the tag is to have stronger discriminative power in gait recognition.

- **Deep Reinforcement Learning on Tag Selection**. As different tag signals reflect motion of user's different body parts (e.g., torso, lower limb), not all the tags we used are of equal importance. We can treat each tag as a feature of time series data and select the feature by tag selection. Since tag selection is a process of selecting the optimal subset of tags, we can learn from the idea of *Reinforcement Learning* to optimize the time series data, which is a Markov decision process (MDP). On this basis, we use *Reinforcement Learning* to progressively select the optimal subset of tags in each iteration. Figure 7 provides the architecture of TSNet. The Agent interacts with the environment and uses the rewards generated by the environment to generate actions and update its state. We use the *Policy Gradient* algorithm to perform tag selection, and adjust the direction of gradient descent by maximizing a set of continuous action discounted reward, finally resulting in a given number m of the optimal tag subset. The states, actions and rewards of this Reinforcement Learning are described as follows.

  States: The state S of the Reinforcement Learning, as an input of agent, consists of two separate parts $\{S_1, S_2\}$. $S_1 = [M, F]$, which is the concatenation of two tensors M and F. F is global information of a sample, which is information composed of 25 tags, and M is information indicating m tags selected from F. M is introduced to implicitly provide TSNet with knowledge of which tags of tag selection. S_2, the binary mask of the selected tag indices, intended to explicitly make the TSNet aware of the selection.

  Actions: The action is the output of the TSNet $TS(S;\theta)$, which is the adjustment direction of each selected tag. In our case, we define 3 types of actions as 'moving to left'(action 0), 'stopping'(action 1) and 'moving to right'(action 2), and setting the moving step to 1. As shown in Figure 7, the TSNet generates a vector $A \in R^{mx3}$ at each iteration, and $A_{i,j} \in [0, 1]$ represents the probability of choosing action j for the ith selected tag. Then, the tag adjustments can be set as:

$$M_i^{'} = M_i + \Delta_i \tag{7}$$

Where Δ_i is defined as:

$$\Delta_i = \begin{cases} -min\{1, (M_i - \eta_i)\} & \text{if action 0} \\ 0 & \text{if action 1} \\ +min\{1, (\gamma_i - M_i - 1)\} & \text{if action 2} \end{cases} \tag{8}$$

In order to ensure that the action is executed correctly, for example the index of the last selected tag should not be larger than f. Therefore, We set the upper boundary $\gamma_i (i = 1, 2, ..., m)$ and the lower boundary $\eta_i (i = 1, 2, ..., m)$ of the tag adjustments, which are the middle between the current tag and the next one and the middle between the current tag and the previous one in the selected tag set, respectively:

$$\gamma = \begin{cases} \lceil (M_i + M_{i+1})/2 \rceil & 1 \le i \le m-1 \\ \text{f} & i = m \end{cases} \tag{9}$$

$$\eta_i = \begin{cases} \lceil (M_i + M_{i-1})/2 \rceil & 2 \le i \le m \\ 0 & i = 1 \end{cases} \tag{10}$$

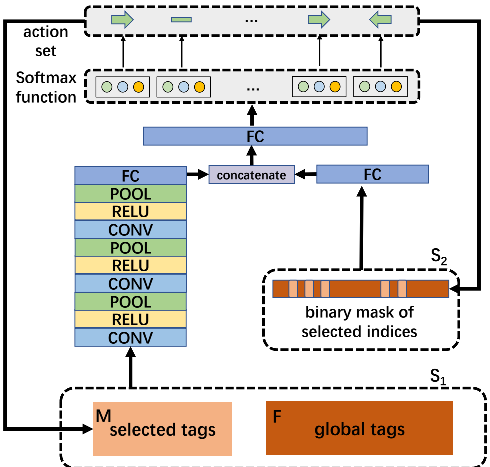


Fig. 7. The architecture of TSNet for deep reinforcement learning on tag selection

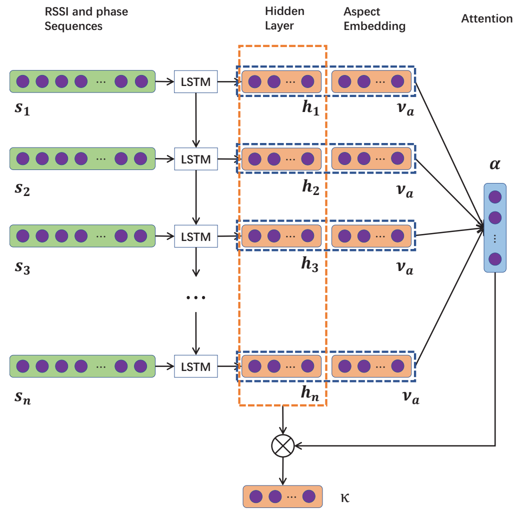


Fig. 8. The architecture of the *AT-LSTM* model.

where '⌈⌉' represents the ceil function, m is the number of selected tags and f is the number of tags. Here γ and η are two arrays with size m. The adjustment of a tag_i is executed within the boundary $[\gamma, \eta_i)$, or otherwise invalid.

Rewards: The rewards indicate how good the agent is after performing an action in state S, which is a function $r(s_t, a_t)$. In our work, we use pre-trained AT-LSTM model to generate rewards to guide Tag Selection. In our reinforcement learning, the output of our pre-trained model is not true or false, but only accuracy. So for the first iteration, the reward r is set to 1 if the test accuracy is above a pre-defined accuracy threshold Λ, -1 otherwise. In order to better represent the change of reward r, we use $r^{'}$ to indicate the difference in the accuracy (P_n and P_{n-1}) between the $(n-1)th$ and nth iteration. We define the $r^{'}$ as follows:

$$r^{'} = P_n - P_{n-1} \tag{11}$$

In Reinforcement Learning, a strong stimulation of $r = +\xi$ is enforced when the $r^{'}$ is larger than a pre-defined threshold Ω, and a strong punishment of $r = -\xi$ if $r^{'}$ is less than $-\Omega$. Otherwise, the reward r takes value in $\{-1, 1\}$. For $nth(n > 1)$ iteration, the reward r can be written as:

$$r = \begin{cases} +\xi & \text{if } r^{'} >= \Omega \\ -\xi & \text{if } r^{'} <= -\Omega \\ \{-1, 1\} & \text{otherwise} \end{cases} \tag{12}$$

Policy Gradient methods for Reinforcement Learning. Figure 7 shows the entire architecture of the TSNet $TS(S;\theta)$, which uses the Convolutional Neural Network (CNN) as a fitting function for Reinforcement Learning. The two separate parts of S are sent into the TSNet separately, as S_1 is fed into 3 convolutional network followed by one fully connected layer and S_2 is fed into a fully connected layer. Then, the outputs of the two fully connected layers are concatenated and fed into the third fully connected layer. Finally, it predicts the optimal action set by the softmax function.

Our goal is to maximize the objective function $J(\theta) = \mathbb{E}(r_1 + \gamma r_2 + \gamma^2 r_3 + \ldots|\pi(S,\theta))$ (the discounted reward), we compute the cross-entropy loss as follows:

$$L(\theta) = -\frac{1}{m}\sum_{i=1}^{m} log\pi(a_t|s_t,\theta)v_t \tag{13}$$

where $\pi(a_t|s_t,\theta)$ is the action probability of the sotfmax function output, and v_t is a normalized discounted reward, which is mainly used to speed up the gradient descent. Our loss is multiplied by v_t in the original cross-entropy form, using v_t to indicate whether the cross-entropy calculated gradient is a trustworthy gradient.If v_t is small or negative, it means that the gradient descent is going in the wrong direction. We should update the parameters in the other direction. If the v_t is positive or large, v_t will praise the gradient of cross-entropy and go down in that direction. The loss provides a direction for the Reinforcement Learning to update parameter θ which is updated as follows:

$$\theta = \theta + \beta\nabla L(\theta) \tag{14}$$

where β is the learning rate of the Deep Reinforcement Learning.

Algorithm 2: DRL on Tag Selection

Input: Training data $\boldsymbol{T}$, the number of iterations $\boldsymbol{it}$, the number of steps in one episode $\boldsymbol{ep}$, AT-LSTM model $\boldsymbol{AT}$

Output: Weight θ of TSNet $\boldsymbol{TS(S;\theta)}$

initialise θ;
for $epoch = 1, 2, \ldots, \boldsymbol{it}$ **do**
    randomly choose m tags from $\boldsymbol{T}$;
    initialise $S = \{S_1, S_2\}$;
    for $t = 1, 2, \ldots, \boldsymbol{ep}$ **do**
        use S_t to generate $A_t = TS(S_t;\theta)$;
        choose the action w.r.t. A_t;
        update the selected tags to M_{t+1} by equation (7);
        update the state to S_{t+1};
        compute the reward r_t using $\boldsymbol{AT}$, by equation (12);
    end
    compute the normalized discounted reward v_t;
    compute the loss $L(\theta)$ by equation (13);
    update θ by equation (14);
end
return θ;

Combining the AT-LSTM and TSNet. For the Deep Reinforcement Learning on Tag Selection, we first select m in f tags, m is pre-set. The selected m tags are used to train TSNet, and the pre-trained AT-LSTM model provides TSNet with real-time accuracy as TSNet's reward. For different m, the global data is re-cleaned according to the m tags selected by the TSNet. These new data are used to train and test the AT-LSTM to obtain the best model and the optimal tag subset. These two models promote each other mutually, and after iterations, the entire Reinforcement Learning precess will converge and tend to be stable. The pipeline of our Deep Reinforcement Learning on Tag Selection is summarized in Algorithm 2.

After running the Relief F algorithm, computing the F-score for each tag and running the TSNet, the correlation of each tag is obtained, and the importance of each tag is sorted according to the F-score. More importantly, we can automatically obtain the optimal tag subset based TSNet. Finally, all three methods can select the $top - N$ tag based on the specific application scenarios.

4.4 Edge Layer: Attention-based LSTM for Accurate Gait Recognition

The edge receives the RSSI and the phase data collected from the selected tags, and performs machine learning algorithms to efficiently recognize human gaits. In this section, we introduce the accurate gait recognition algorithm used in the system.

4.4.1 LSTM with Attention Mechanism. *LSTM*[13] is a popular recurrent neural network model that has been recently largely used and provide excellent performance in the sequence data matching and classification *LSTM*. Since the gait recognition in this system uses the sequential RSSI and phase signal as input, it can be naturally modeled as a time sequence classification problem. As a result, we adopt the LSTM model in this work for gait training and classification to achieve accurate gait recognition in smart spaces. To achieve high recognition accuracy, the attention mechanism is further introduced to enhance the basic LSTM network [16]. Figure 8 illustrates the network structure of our attention-based LSTM model.

Let $\{(r_1^1, p_1^1,\ r_1^2, p_1^2,\ \cdots, r_1^k, p_1^k),\ (r_2^1, p_2^1,\ r_2^2, p_2^2,\ \cdots, r_2^k, p_2^k), \cdots, (r_n^1, p_n^1,\ r_n^2, p_n^2,\ \cdots, r_n^k, p_n^k)\}$ be the RSSI strength and phase sequences of the k selected tags with the max length of n, those sequences whose length is shorter than n will be zero-padded. $\{h_1, h_2, \cdots, h_n\}$ represents the hidden layer. v_a is the aspect embedding, and $\{\alpha_1, \alpha_2, \cdots, \alpha_n\}$ indicates the attention layer. The following operators are defined [34]:

$$\kappa = \tanh(\begin{bmatrix} W_H H \\ W_{v_a}(v_a e_n) \end{bmatrix}) \tag{15}$$

$$\alpha = softmax(w^T \kappa) \tag{16}$$

$$\varphi = H\alpha^T \tag{17}$$

where $H \in \mathbb{R}^{d\times n}$, $e_n \in \mathbb{R}^{1\times n}$, $v_a \in \mathbb{R}^{da}$, $W_{v_a} \in \mathbb{R}^{d_a\times d_a}$, $W_H \in \mathbb{R}^{d\times d}$, $\kappa \in \mathbb{R}^{(d+d_a)\times n}$, $w \in \mathbb{R}^{d+d_a}$, $\alpha \in \mathbb{R}^{1\times n}$, $\varphi \in \mathbb{R}^d$ are the attention-based LSTM parameters and d is the number of neurons of the LSTM. d_a is the aspect embedding dimension parameter, and e_n is a row vector with n 1s. The $v_a e_n = [v_a; v_a; ...; v_a]$ therefore concatenates v_a for n times. And the gait representation of each user is then given by:

$$h^{proj} = \tanh(W_\varphi \varphi + W_{h_n} h_n) \tag{18}$$

where $W_\varphi \in \mathbb{R}^{d\times d}$, $W_{h_n} \in \mathbb{R}^{d\times d}$, $h^{proj} \in \mathbb{R}^{d\times d}$. h^{proj} is considered as a feature representation of the RSSI strength and phase sequences. The classification of the gait is given by:

$$\hat{y} = softmax(W_{h^{proj}} h^{proj} + b_{h^{proj}}) \tag{19}$$

where $W_{h^{proj}} \in \mathbb{R}^{c\times d}$, $b_{h^{proj}} \in \mathbb{R}^c$ are the parameters of the *softmax* layer, c is the number of class and is the total number of users in the system. The loss function is given by [34]:

$$loss = -\sum_i \sum_j y_i^j log \hat{y}_i^j + \lambda \|\theta\|^2 \tag{20}$$

where i is the index of the i-th RSSI strength and phase sequences of the k selected tags, j is the index of the j-th class. y represents the one-hot ground truth and $\hat{y}$ is the probability prediction. The θ represents the parameter setting of the model and λ is the L_2-Regularization term [34]. The attention-based LSTM model is then trained by minimizing the above loss function.

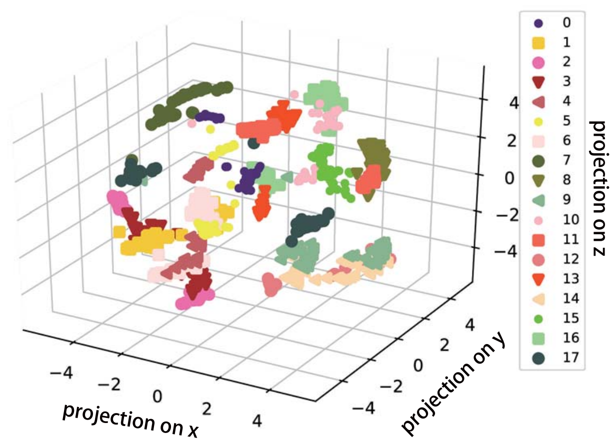


Fig. 9. t-SNE projection

To visually inspect the training performance, we perform t-SNE projection [17] to project all gait features into the 3D space. Figure 9 shows the t-SNE projection [17] of the learned features for 18 users. As shown in the figure, the projected features can be visually identified as different disjoint clusters in the 3D space, indicating that the attention-based LSTM model can be used to effectively recognition gaits based on the captured RSSI and phase information.

4.5 Service Provider Layer and Application Layer

With the device layer and the edge layer, the gait recognition functionality is achieved. To provide gait recognition services for user identification in smart spaces, the service needs to be distributed to the customers. Therefore, following the GRaaS architecture, the gait recognition services is registered by the edge to the service provider layer, which then serves as a central cloud platform for service management. The customers in the application layer retrieves the service from the cloud, and then is able to achieve user identification using gait to support applications such as authentication in the smart buildings, and user tracking in smart cities, etc.

5 EVALUATIONS

In this section, we provide detailed evaluations on the RFID-based gait recognition system. We first briefly describe the experimental settings, and then summarize the detailed results, including the evaluations on the data source, tag selection, the final gait recognition performance, with a comparison between our attention-based LSTM model and traditional classification algorithms such as Sparse Representation Classification (SRC) [26], and the energy efficiency by adopting the GRaaS model.

5.1 Experiment Settings

5.1.1 Implementation. We implement the prototype of the RFID-based gait recognition system using commercially off-the-shelf devices. The system prototype includes a commercial RFID reader *Impinj R420* and off-the-shelf UHF passive tag *Alien az-9662*[14], where the directional antenna of *Impinj* reader is *Laird A9028PCL*, with frequency range 902-928mHz, and the gain 9dBi. As illustrated in Figure 5, the whole system is deployed in a corridor. Each tag is separated by 10 centimeters to form a 5×5 tag matrix.The antenna is connected to the reader, and the reader is connected to the router through the network cable. We use a laptop as the edge device to collect the data from the RFID reader and perform gait recognition.

5.1.2 Dataset. To evaluate the system performance, we recruit 18 participants to walk freely along the corridor for 15 minutes, and we have 900 instances totally. We use half of the data as the training data and the rest as the testing dataset.

5.2 Evaluation Results

5.2.1 Impact of data source. To understand the impact of different data sources, i.e., RSSI, phase, and the combination of these two on the recognition accuracy, we measure the gait recognition accuracies by varying data sources and different classification algorithms. As shown in Figure 10, for both SRC and attention-based LSTM, the data sources has observable impacts on the final recognition accuracy. For SRC, using RSSI, phase, and the combination of two achieves 75.1%, 64%, and 74.8% accuracy. Using attention-based LSTM, on the other hand, the accuracy rises to 90.2%, 92.8%, and 96.3% respectively. This dues to the fact that more information the data contains, the richer its features, and LSTM is able to learn characteristics of gait more effectively, thus combining both RSSI and phase improves the classification accuracy. Therefore, both RSSI and phase are used for gait recognition in the attention-based LSTM model.

5.2.2 Impact of tag selection. Selecting the subset of tags as input not only reduces the computational cost, but also provides a more robust gait recognition performance. Figure 11 and Figure 12 show the gait recognition accuracy without tag selection and by using the subset of tags selected by the Relief F and F-Statistics algorithms respectively. As shown in Figure 11, if the tags are randomly selected from 25 tag matrix, the accuracy increases when more tags are used but increases slower than using the tag selection algorithms shown in Figure 12. As shown in Figure 12, for both SRC and attention-based LSTM, the recognition accuracy increases as the number of selected tags increases and is faster than no tag selections. This is due to the fact that more tags are used the trained models using the RSSI and phase better reflect the gait movements of the whole body, and results in higher recognition accuracy, and with tag selection, smaller number of tags are required to achieve the same recognition performance. As shown in Figure 12, for SRC using Relief F algorithm, the accuracy of using 15 tags is comparable to that of using 25 tags. This significantly reduces the training and recognition costs and speeds up the recognition process, which is especially useful for IoT applications. For attention-based LSTM, similar observations can be made, by using the Relief F algorithms only 14 tags and F-statistics only 9 tags are used to achieve an accuracy of 96%, which is comparable to the accuracy of using all 25 tags. Likewise, the Deep Reinforcement Learning method shows similar pattern. The results show that the three tag selection algorithms can be used to effectively select a subset of tags to achieve high recognition accuracy while significantly reducing the computational cost. More importantly, the Deep Reinforcement Learning on tag selection is more stable than other two methods, which means it is more robust.

5.2.3 Impact of number of users. To understand the system scalability, we further study the RFID-based gait recognition performance as the number of users changes in the system. As shown in Figure 13, we analyzes the recognition accuracy by varying the number of users from 2 to 18. As can be seen from the figure that for SRC, as the number of samples increases, the classification accuracy of SRC gradually decreases, which shows that SRC does not scale well and its robustness is limited as the number of people increases. For attention-based LSTM, the classification accuracy does not directly affected by user number, but maintains a relatively stable classification accuracy at around 96.3%, indicating that its performance is robust in multi-user environments and can be scaled for multiple people classification using RFID-based gait recognition.

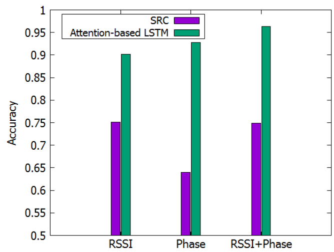


Fig. 10. Impact of data sources

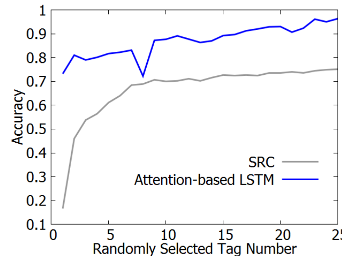


Fig. 11. Gait recognition performance without tag selection

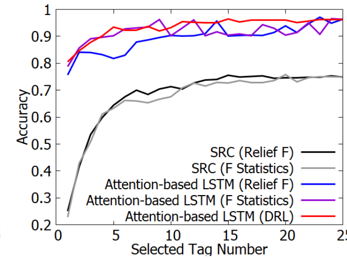


Fig. 12. Impact of tag selection on recognition accuracy

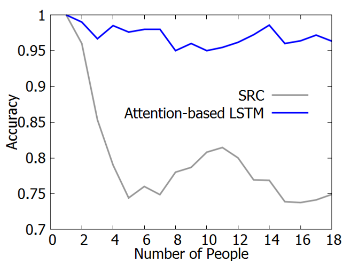


Fig. 13. Impact of user number on the recognition accuracy

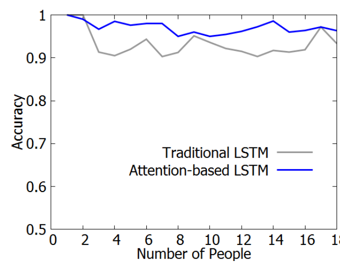


Fig. 14. Effectiveness of attention mechanism

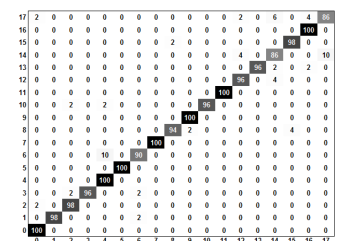


Fig. 15. Confusion matrix of RFID-based gait recognition

5.2.4 Impact of attention model. The attention mechanism used in this work aims at improving the accuracy and robustness in the gait recognition. To understand its effectiveness, we compared the performance of traditional LSTM network and the attention-based LSTM used in this work. As shown in Figure 14, it can be seen that the gait recognition accuracy of traditional LSTM model gradually reduced as the number of users increases, showing its degraded performance and scalability when user number becomes large. On the other hand, the classification accuracy of attention-based LSTM remains high and almost unaffected by the increasing number of users, indicating that it is more robust, has more stable performance, and better scalability. Figure 15 shows the confusion matrix for all 18 users, and as can be seen from the figure that all users can be efficiently recognized using the RSSI and phase information of RFID, on average the gait recognition accuracy is 96.3%.

5.2.5 Energy Analysis. The GRaaS architecture proposed in this work adopts the emerging edge computing paradigm and allows computation in the IoT devices to be offloaded to the edge for better energy efficiency. In this work, the device layer is mainly responsible for deciding the subset of tags and collecting RSSI and phase information from the tags. The key gait recognition algorithms which constitute the major computation cost are offloaded to edge device which has much stronger hardware capability for more efficient computation. As shown in Table 2, we measure the power consumption for both local computation and data offloading. As shown in the table, computing the attention-based LSTM for gait recognition consumes 472.5 mJ energy, while offloading to the edge server only consumes 48 mJ. Since the edge device usually is not power-constraint and has stronger computational power, adopting the GRaaS model for gait recognition service generation significantly reduces the energy consumption for the IoT devices in the device layer, which makes the GRaaS model especially suitable for IoT applications.

Table 2. Computation and communication overhead of RFID-based gait recognition in GRaaS

Task	Time (ms)	Energy (mJ)
Attention-based LSTM	315	472.5
Data Transmission	8	48

6 DISCUSSIONS

6.1 Applications

The proposed GRaaS framework and the RFID-based gait recognition services has great potential in smart space applications and can act as a service in smart homes, smart buildings, smart cities, etc. Following are two possible use scenario examples: (1) Smart Homes. In smart home applications, the user identification is essential to provide customized services based on different family member's living hobbies. (2) Smart Offices. In smart office, authentication is one critical service that needs to be realized. With the RFID-based gait recognition, user's can be accurately identified and authenticated without needing to use a key or identity card. We envision there is great potential for the GRaaS framework and the RFID-based gait recognition services in future smart space applications.

6.2 Limitations and Future Work

Although the proposed RFID-based gait recognition service can achieve high-precision recognition for large groups, the granularity of the current recognized gait is still relatively coarse. Secondly, in dynamic environments, the environmental noises will affect the recognition accuracy. Although our learning algorithms can adapt to different environments given enough training data is provided, acquiring such high-quality training data for different environments still is a challenges and we leave it as a future work to explore. In addition, when multiple user are walking together, the wireless signals will be affected by all users, and more sophisticated signal separation techniques are required for more robust recognition. Given these limitations, since RFID devices are easier to deploy and cover a wider range, the non-intrusiveness of this approach still make it an attractive alternative approach to traditional identification methods using cameras, acoustic signals, etc. By incorporating with other wireless signals such as WiFi, it might achieve more robust recognition performance in dynamic environments. In the future work, more robust gait recognition services with multi-modal wireless signal input in dynamic environments can be explored for multi-person simultaneous recognition. To adapt to new environments faster and achieve better scalability, techniques such as transfer learning [21] might be possible to used in the system for faster system re-deployment with lesser re-training overhead.

7 CONCLUSIONS

In this work, we propose a Gait Recognition as a Service (GRaaS) model, an instantiation of the popular sensing as a service model for design and implementation of gait recognition services in smart spaces. As a case study, we propose a RFID-based gait recognition service following the GRaaS model. Tags selection algorithms are used and attention-based LSTM model is designed for accurate RFID-based gait recognition, which achieves 96.3% accuracy. Due to its cost-effectiveness, high accuracy, and non-obtrusiveness, the proposed framework and RFID-based gait recognition services have great potential in future smart spaces applications.

ACKNOWLEDGMENTS

This research is supported by the National Natural Science Foundation of China (Grant No. 61972263, 61602319, U1713212), and Natural Science Foundation of Guangdong Province (Grant No. 2017A030313357).

REFERENCES

[1] Mohammed Al-khafajiy, Thar Baker, Carl Chalmers, Muhammad Asim, Hoshang Kolivand, Muhammad Fahim, and Atif Waraich. 2019. Remote health monitoring of elderly through wearable sensors. *Multimedia Tools and Applications* (2019), 1–26.

[2] Mohammed Al-Khafajiy, Lee Webster, Thar Baker, and Atif Waraich. 2018. Towards fog driven IoT healthcare: challenges and framework of fog computing in healthcare. In *Proceedings of the 2nd International Conference on Future Networks and Distributed Systems*. ACM, 9.

[3] Prith Banerjee, Richard Friedrich, Cullen Bash, Patrick Goldsack, Bernardo Huberman, John Manley, Chandrakant Patel, Parthasarathy Ranganathan, and Alistair Veitch. 2011. Everything as a service: Powering the new information economy. *Computer* 44, 3 (2011), 36–43.

[4] Kai Cao and Anil K Jain. 2018. Automated latent fingerprint recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2018).

[5] Dan Chen, Yangyang Hu, Lizhe Wang, Albert Y Zomaya, and Xiaoli Li. 2016. H-PARAFAC: Hierarchical parallel factor analysis of multidimensional big data. *IEEE Transactions on Parallel and Distributed Systems* 28, 4 (2016), 1091–1104.

[6] Dan Chen, Yunbo Tang, Hao Zhang, Lizhe Wang, and Xiaoli Li. 2019. Incremental factorization of big time series data with blind factor approximation. *IEEE Transactions on Knowledge and Data Engineering* (2019).

[7] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N Padmanabhan. 2010. Indoor localization without the pain. In *Proceedings of the sixteenth annual international conference on Mobile computing and networking (MobiCom)*. ACM, 173–184.

[8] Mohammed Dighriri, Gyu Myoung Lee, and Thar Baker. 2018. Big data environment for smart healthcare applications over 5g mobile network. In *Applications of Big Data Analytics*. Springer, 1–29.

[9] Yuan Gao, Jiayi Ma, and Alan L Yuille. 2017. Semi-supervised sparse representation based classification for face recognition with insufficient labeled samples. *IEEE Transactions on Image Processing* 26, 5 (2017), 2545–2560.

[10] Brij B Gupta and Quan Z Sheng. 2019. *Machine Learning for Computer and Cyber Security: Principle, Algorithms, and Practices*. CRC Press.

[11] Ju Han and Bir Bhanu. 2004. Statistical feature fusion for gait-based human recognition. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2. IEEE.

[12] Jinsong Han, Wei Xi, Kun Zhao, and Zhiping Jiang. 2014. Device-Free Object Tracking Using Passive Tags. In *INFOCOM, 2014 Proceedings IEEE*. 1605–1617.

[13] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-term Memory. 9 (12 1997), 1735–80.

[14] Impinj. 2013. Speedway revolution reader application note - low level user data support.

[15] Hengjin Ke, Dan Chen, Tejal Shah, Xianzeng Liu, Xinhua Zhang, Lei Zhang, and Xiaoli Li. 2018. Cloud-aided online EEG classification system for brain healthcare: A case study of depression evaluation with a lightweight CNN. *Software: Practice and Experience* (2018).

[16] Jun Liu, Gang Wang, Ping Hu, Ling Yu Duan, and Alex C. Kot. 2017. Global Context-Aware Attention LSTM Networks for 3D Action Recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*. 3671–3680.

[17] Laurens Van Der Maaten and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 2605 (2008), 2579–2605.

[18] Pattabhi Ramaiah Nalla and Ajay Kumar. 2017. Toward more accurate iris recognition using cross-spectral matching. *IEEE Transactions on Image processing* 26, 1 (2017), 208–221.

[19] Mark S. Nixon and John N. Carter. 2007. Automatic Recognition by Gait. *Proc. IEEE* 94, 11 (2007), 2013–2024.

[20] Soraia Oueida, Yehia Kotb, Moayad Aloqaily, Yaser Jararweh, and Thar Baker. 2018. An edge computing based smart healthcare framework for resource management. *Sensors* 18, 12 (2018), 4307.

[21] Sinno Jialin Pan and Qiang Yang. 2009. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering* 22, 10 (2009), 1345–1359.

[22] Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos. 2014. Sensing as a service model for smart cities supported by internet of things. *Transactions on Emerging Telecommunications Technologies* 25, 1 (2014), 81–93.

[23] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *International Conference on Mobile Computing & NETWORKING*. 27–38.

[24] Douglas A Reynolds. 1995. Speaker identification and verification using Gaussian mixture speaker models. *Speech communication* 17, 1-2 (1995), 91–108.

[25] Longfei Shangguan, Zheng Yang, Alex X Liu, Zimu Zhou, and Yunhao Liu. 2015. Relative Localization of RFID Tags using Spatial-Temporal Phase Profiling.. In *NSDI*. 251–263.

[26] Yiran Shen, Wen Hu, Mingrui Yang, Bo Wei, Simon Lucey, and Chun Tung Chou. 2014. Face recognition on smartphones via optimised sparse representation classification. In *Proceedings of the 13th international symposium on Information processing in sensor networks (IPSN)*. IEEE Press, 237–248.

[27] Yunbo Tang, Dan Chen, Lizhe Wang, Albert Y Zomaya, Jingying Chen, and Honghai Liu. 2018. Bayesian tensor factorization for multi-way analysis of multi-dimensional EEG. *Neurocomputing* 318 (2018), 162–174.
[28] Stephan Wagner, Marcus Handte, Marco Zuniga, and Pedro José Marrón. 2012. On optimal tag placement for indoor localization. In *IEEE International Conference on Pervasive Computing and Communications*. 162–170.
[29] Chen Wang, Junping Zhang, Liang Wang, Jian Pu, and Xiaoru Yuan. 2012. Human Identification Using Temporal Information Preserving Gait Template. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34, 11 (2012), 2164–2176.
[30] Liang Wang, Tieniu Tan, Huazhong Ning, and Weiming Hu. 2003. Silhouette analysis-based gait recognition for human identification. *Pattern Analysis and Machine Intelligence IEEE Transactions on* 25, 12 (2003), 1505–1518.
[31] Wei Wang, Alex X. Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 363–373.
[32] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking (MobiCom)*. ACM, 65–76.
[33] Xiaohang Wang, Jin Song Dong, Chung-Yau Chin, SankaRavipriya Hettiarachchi, and Daqing Zhang. 2004. Semantic space: An infrastructure for smart spaces. *IEEE Pervasive computing* 3, 3 (2004), 32–39.
[34] Yequan Wang, Minlie Huang, Li Zhao, et al. 2016. Attention-based LSTM for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing*. 606–615.
[35] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking (MobiCom)*. ACM, 617–628.
[36] Weitao Xu, Chitra Javali, Girish Revadigar, Chengwen Luo, Neil Bergmann, and Wen Hu. 2017. Gait-key: A gait-based shared secret key generation protocol for wearable devices. *ACM Transactions on Sensor Networks (TOSN)* 13, 1 (2017), 6.
[37] Lina Yao, Quan Z Sheng, Wenjie Ruan, Tao Gu, Xue Li, Nick Falkner, and Zhi Yang. 2015. RF-Care: Device-Free Posture Recognition for Elderly People Using A Passive RFID Tag Array. In *EAI Endorsed Transactions*.
[38] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere. 2016. WiFi-ID: Human Identification Using WiFi Signal. In *2016 International Conference on Distributed Computing in Sensor Systems (DCOSS)*. 75–82.
[39] Jin Zhang, Bo Wei, Wen Hu, and Salil S. Kanhere. 2016. WiFi-ID: Human Identification Using WiFi Signal. In *International Conference on Distributed Computing in Sensor Systems*. 75–82.
[40] Yongpan Zou, Jiang Xiao, Jinsong Han, Kaishun Wu, Yun Li, and Lionel M. Ni. 2017. GRfid: A Device-Free RFID-Based Gesture Recognition System. *IEEE Transactions on Mobile Computing* 16, 2 (2017), 381–393.